DMOS DEVICE OF SMALL DIMENSIONS AND MANUFACTURING PROCESS THEREOF

PRIORITY CLAIM

[1] This application claims priority from Italian patent application No. TO2003A000013, filed January 14, 2003, which is incorporated herein by reference.

TECHNICAL FIELD

[2] The present invention relates to a DMOS device of small dimensions and the manufacturing process thereof. In particular, the invention relates to a DMOS transistor in BCD (Bipolar CMOS and DMOS) technology capable of operating at voltages higher than 16 V.

BACKGROUND

[3] As is known, BCD power technology enables integration of structures of different types in a same chip. This technology has enjoyed wide application thanks to integration of circuitry transistors defining an intelligent part with power components.

[4] Consequently, in view of the continuous requirements of miniaturization, it is desirable to modify the present process flow, so as to reduce the dimensions of the devices, and specifically of the DMOS transistors.

[5] In particular, it is desirable to reduce the size between the source contact and the gate region of the DMOS transistor, without causing at the same time any critical factors in the performance of the device or of the fabrication process.

[6] On the other hand, a mere reduction of the dimensions and distances between the various parts without modifying the layout of the device would entail the risk of errors in the positioning of the various regions or superposition thereof on account of the tolerances of fabrication, and hence of malfunctioning of the device.

[7] The aim of the present invention is to solve the problems referred to above.

SUMMARY

[8] According to an aspect of the present invention, a DMOS device and the corresponding manufacturing process are provided, as defined in claim 1 and 6, respectively.

[9] In practice, according to one aspect of the invention, the enriched contact regions, necessary for contacting the body region formed in the source active area, are formed after opening the contacts, in a self-aligned manner to the contacts themselves (self-aligned body contact implant). In this way, the body contact implant is performed only where it is necessary to obtain contact with the body region; consequently, there is a gain in tolerance, and it is possible to reduce the distance between the body contact and the gate region and hence the size of the source active area, without giving rise to any critical factors.

BRIEF DESCRIPTION OF THE DRAWINGS

[10] For a better understanding of the invention, embodiments thereof are now described, purely by way of non-limiting example, with reference to the attached drawings, wherein:

[11] **FIG. 1** illustrates a mask used in a first fabrication step of a known DMOS device;

[12] **FIG. 2** illustrates a cross-section through a portion of a wafer, taken along section line II-II of **FIG. 1**, after the step using the mask of **FIG. 1**;

[13] **FIGS. 3 and 4** illustrate a mask used in a fabrication step subsequent to the step of **FIG. 1** and the cross-section thereof, taken in a plane similar to that of **FIG. 2**;

[14] **FIG. 5** illustrates a cross-section obtained in a subsequent fabrication step;

[15] **FIGS. 6 and 7** illustrate a mask used in a fabrication step subsequent to the step of **FIG. 5** and the cross-section thereof;

[16] **FIGS. 8 and 9** illustrate a mask used in a fabrication step subsequent to the step of **FIG. 6** and the cross-section thereof;

[17] **FIG. 10** illustrates a longitudinal cross-section of a portion of the wafer, taken along section line X-X of **FIG. 9**;

[18] **FIGS. 11-13** illustrate a mask used in a fabrication step subsequent to the step of **FIG. 8** and the cross-section thereof;

[19] **FIGS. 14 and 15** illustrate a mask used in a final fabrication step and the cross-section of a known DMOS device;

[20] **FIG. 16** illustrates masks used according to a first embodiment of the invention;

[21] **FIG. 17** illustrates a cross-section taken along section line XVII-XVII of **FIG. 16**;

[22] **FIG. 18** illustrates a mask used in a subsequent fabrication step according to one embodiment of the present invention;

[23] **FIGS. 19 and 20** illustrate two cross-sections taken along the section planes XIX-XIX and XX-XX of **FIG. 18**;

[24] **FIG. 21** illustrates a mask used in a subsequent fabrication step according to one embodiment of the present invention;

[25] **FIG. 22** illustrates a cross-section taken along the section plane XXII-XXII of **FIG. 21**;

[26] **FIG. 23** illustrates a cross-section similar to that of **FIG. 19**, taken in a subsequent fabrication step according to one embodiment of the present invention;

[27] **FIG. 24** illustrates a mask used in a subsequent fabrication step according to one embodiment of the present invention;

[28] **FIG. 25** illustrates a cross-section taken along section plane XXV-XXV, similar to that of **FIG. 23**;

[29] **FIG. 26** illustrates a cross-section similar to **FIG. 25**, in a subsequent fabrication step according to one embodiment of the present invention;

[30] **FIG. 27** illustrates a cross-section similar to **FIG. 20**, in a subsequent fabrication step according to one embodiment of the present invention;

[31] **FIG. 28** illustrates a mask used in a fabrication process alternative to that of **FIG. 16** according to another embodiment of the present invention; and

[32] **FIG. 29** illustrates a cross-section taken along section plane XXIX-XXIX.

DETAILED DESCRIPTION

[33] The following discussion is presented to enable a person skilled in the art to make and use the invention. Various modifications to the embodiments will be readily apparent to those skilled in the art, and the generic principles herein may be applied to other embodiments and applications without departing from the spirit and scope of the present invention. Thus, the present invention is not intended to be limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features disclosed herein.

[34] Initially, a known process will be described, in order to highlight the critical aspects thereof and enable detection of the differences with respect to the described embodiments of the present invention.

[35] As illustrated in **FIGS. 1 and 2**, initially, in a wafer ***1*** of semiconductor material, comprising a standard-doping substrate ***4***, accommodating at least one well ***2***, here of N type, and a surface ***8***, an active-area mask ***5*** is formed, that has the aim of protecting the areas of the substrate intended to accommodate the conductive regions forming the devices to be integrated (here a DMOS transistor). In the example, in which the DMOS device must withstand voltages higher than 16 V, the drain and source regions of the DMOS transistor must be formed in separate active areas; consequently, the active-area mask ***5*** has a central region 5a of larger size, defining a source active area, and two lateral regions ***5b*** of smaller size, formed on the two sides of the central region ***5a***, defining the drain active areas. In **FIG. 1**, as in the subsequent figures, the areas coated with the mask are hatched.

[36] The process for defining the active areas is standard and hence not illustrated in detail herein. At the end of the process, as may be seen in **FIG. 2**, a field-oxide region ***3*** extends on the surface (and in part inside) the wafer ***1***, delimiting on all the sides a source active area ***6*** and two drain active areas ***7***. As

may be noted in particular in **FIG. 1**, the source active area ***6*** and drain active areas ***7*** have a rectangular shape, and the two drain active areas ***7*** each extend alongside a respective side of the source active area ***6***.

[37] In a way not illustrated herein, a gate-oxide layer is deposited. Then a polysilicon layer is deposited and defined using a poly mask ***10***, as illustrated in **FIGS. 3 and 4**. After removing the exposed portions, a gate region ***11*** is formed, which has the same shape as the poly mask ***10*** of **FIG. 3** and thus extends along the perimeter of the source area ***6*** and in part on top of the field oxide ***3***, and therefore substantially along the perimeter of a rectangle.

[38] Next, a body mask, not illustrated, is deposited and has an opening substantially matching with the source active area ***6***, and, using this mask, a body region ***12***, of P type, is implanted. At the end of the implantation, the structure of **FIG. 5** is obtained, in which the body region is designated by ***15***.

[39] Subsequently (**FIG. 6**), after oxidation (not described in detail herein), a low-doped drain (LDD) mask ***16*** is deposited which exposes the body region ***15*** except for one or more isolation islands ***17*** and two longitudinal end areas ***18***, which should accommodate body contacts. Using the LDD mask ***16***, an LDD implant is performed, here of N type, so that an LDD region ***19*** is formed inside the body region ***15*** and surrounds one or more non-implanted central portions ***20***, where the body region ***15*** emerges at the surface ***8*** of the wafer ***1***.

[40] Next, **FIGS. 8-10**, in a per se known manner, spacers ***24*** are formed at the sides of the gate region ***11*** (**FIG. 9**), and an S/D mask ***25*** is deposited which, inside the source active area ***6***, has a shape similar to that of the LDD mask ***16***, with islands ***17a*** and longitudinal end areas ***18a***. Subsequently, using the S/D mask ***25***, dopant species of N type are implanted, which, in the drain active areas ***7***, form drain regions ***26*** of N+ type, and, in the source active area ***6***, form a rectangular source region ***27***, narrower than the LDD region ***19*** because of the spacers ***24***. Consequently, the source region ***27*** is surrounded on the two long sides by a peripheral LDD portion ***19*** and surrounds the non-implanted central portion or portions ***20***, where the body region ***15*** extends up to the wafer surface. Then the structure of **FIGS. 9 and 10** is obtained, which illustrate two perpendicular cross-

sections showing a non-implanted central portion ***20*** and the two end regions ***21***, where the body region ***15*** extends up to the surface of the wafer ***1***.

[41] Next (**FIGS. 11-13**), a body-contact mask ***30*** is formed and covers completely the drain active areas ***7*** and a fair part of the source active area ***6***, except for portions where contact regions for the body region ***15*** are to be formed. For this purpose, on top of the source active area ***6***, the body-contact mask ***30*** has a substantially complementary shape to the S/D mask ***25***, except for tolerances. Where previously the non-implanted central portion ***20*** and the end regions ***21*** were present, now body-contact regions ***31***, of P+ type, are formed.

[42] Next, the wafer ***1*** is coated with an insulating layer ***35*** (**FIG. 15**), and the contacts are opened, using a contact mask ***36*** (**FIG. 14**). In particular, in the insulating layer ***35*** there are formed: openings ***37a***, which reach the drain regions ***26*** in the drain active areas ***7***; openings ***37b***, ***37c***, which reach the source region ***27*** and the contact regions ***31*** in the source active area ***6***; and openings ***37d***, which reach the gate regions ***11*** (**FIG. 14**). The openings ***37a-37c*** are then filled with metal material so as to form contacts ***38***, in a per se known manner.

[43] If the aim is to reduce the dimensions of the DMOS transistor of **FIG. 15**, and leave the proportions unaltered, it is possible to act on the dimensions of the contacts ***37***, on the width of the drain active areas ***7***, and on the distance between the contacts ***38*** and the gate region ***11***.

[44] As may be noted in particular from the cross-section of **FIG. 15**, the latter parameter (contacts/gate distance) is somewhat critical. In fact, a possible misalignment and/or a dimensional variation of the S/D mask ***25*** could prevent the source region ***27*** from being made in a correct way in the area comprised between the body-contact region ***31*** and the gate region, since, when the S/D mask ***25*** is made, the spacers ***24*** are already present. In particular (see **FIG. 9**), a misalignment of the S/D mask ***25***, for example its displacement to the right, in the case of reduction of the distance referred to above, would entail the risk of not performing a proper implantation of the portion of the source region ***27*** formed to the right of the contact ***38***. This is all the more serious in consideration of the fact that the area not implanted with N-type species will certainly be P^+-implanted, since

the body-contact mask ***30*** (**FIG. 12**) has an opening wider than the island ***25***, with the risk of pinch-off of the DMOS transistor in this area.

[45] To solve the above problem, according to one aspect of the invention, it is proposed to implant the body contacts after opening the contacts. In this way the P+ implant is performed only where it is necessary, enabling a gain in tolerance and hence a reduction of the contact/body distance. In particular, it is possible to reduce the dimensions of the source active area, without involving critical aspects.

[46] Hereinafter, the differences in the process flow will be described in detail, as compared to the known process described above, according to two embodiments of the invention. Consequently, in **FIGS. 16-29**, the masks and the regions, which are not modified substantially with respect to the known device, are designated by the same reference numbers.

[47] The process starts with the steps described with reference to **FIGS. 1-4**, including defining the active areas ***6***, ***7***, depositing the gate oxide, forming the gate regions ***11***, and forming the body region ***15***. Next, using an LDD mask ***45*** illustrated in **FIG. 16**, the LDD implant is performed, here of N type. In practice, as regards most of the source active area ***6***, the implant is blanket type, i.e., not shielded. The implant then forms, in the source active area ***6***, a well-shaped LDD source region ***50***, which extends almost throughout the length of the source active area ***6***, except for longitudinal end portions, as illustrated in cross-section in **FIG. 17**.

[48] Next (**FIGS. 18-20**), the spacers ***24*** are formed, an S/D mask ***52*** is deposited, and the source/drain implant of N type is performed. The S/D mask ***52*** completely exposes the drain active areas ***7*** and, above the source active area ***6***, covers the longitudinal ends of the source active area ***6***. In addition, in an intermediate region of the source active area ***6***, the mask ***52*** forms islands ***53***, which extend width-wise (perpendicular to the longitudinal direction of the source active area ***6***). As may be seen in particular from **FIG. 19**, the islands ***53*** extend as far as above the two opposite sides of the gate region so that the area visible in the cross-section of **FIG. 19** is not S/D implanted with N dopants, and in the area only the light LDD implant is present (transverse LDD source portion 50'). Instead, the areas upstream and downstream of each island ***53***, proceeding in a longitudinal

direction (as illustrated in **FIG. 20**), are N-implanted and form source regions ***54***. On the sides of the source regions ***54***, longitudinal portions 50" are present. In practice, the source regions ***54*** are separate from one another in a longitudinal direction and are connected electrically at the islands ***53*** through the remaining LDD source regions 50'. In the drain active areas ***7***, drain regions ***55*** are formed, here of N+ type.

[49] Subsequently, a first body-contact mask ***58*** is deposited, and exposes only the longitudinal ends of the body region ***15***, and a P+ implant is performed, as illustrated in **FIGS. 21, 22**. End contact regions ***59***, of P type, are then formed, as may be seen from **FIG. 22**, which also illustrates in part the succession of regions ***50'*** and ***54*** in a longitudinal direction.

[50] Next, **FIG. 23**, the wafer ***1*** is coated with an insulating layer ***35***, and the contacts are opened, analogously to the known device, thus forming openings ***37a*** for the drain regions ***26***, openings ***37b*** for the source regions ***27***, openings ***37c*** for the body-contact region, and openings ***37d*** for the gate regions ***11*** (of which only the openings ***37a*** and ***37c*** may be seen in **FIG. 23**).

[51] Next, **FIGS. 24 and 25**, a second body-contact mask ***60*** is deposited, and an implant of P+ type is performed, referred to as open-contact implant. In one embodiment, the P+ implant includes a first, deep implantation step such as to ensure that the body region ***15*** is reached through the transverse portions ***50'***, for example using B^{11}, with an energy of approximately 35 keV and a dose of $5*10^{13}$, and a second, superficial implantation step, for example using BF_2, with an energy of approximately 40 keV and a dose of $5*10^{14}$. Then body-contact regions ***61*** are formed.

[52] Finally, the openings ***37a-37d*** are filled with metal, so as to form the contacts ***38***, as illustrated in the cross-section of **FIGS. 26 and 27**, taken along two parallel planes.

[53] In practice, at the end of the process, the more doped source region ***54*** is formed by a series of portions (here three) that are separate from one another, and the LDD source region is formed by two peripheral portions ***50"*** and by two transverse portions ***50'***. The peripheral portions ***50"*** of the LDD source region ***50***

(the cross section whereof may be seen in **FIG. 27**) extend parallel to the longitudinal direction of the active areas ***6***, ***7*** and each face a respective drain region ***55***. The transverse portions ***50'*** (one of which may be seen in **FIG. 26**) separate physically and connect electrically the second implanted regions ***54*** and are interrupted centrally by the body-contact region ***61***.

[54] According to a different embodiment of the process just described, the LDD implant uses a mask, which, at the source active area ***6***, covers the portions of the body region ***15***, where the body contacts ***61*** must be formed. In practice, as illustrated in **FIG. 28**, the LDD mask, designated by ***45'***, has islands ***64*** that are narrower than the islands ***53*** of the S/D mask ***52*** of **FIG. 18**, to enable LDD implantation on all sides of the body-contact regions. In this way, as may be seen in the cross-section of **FIG. 29**, underneath the islands ***64*** non-implanted central portions ***65*** of P type are present, where the body region ***15*** emerges at the surface of the wafer ***1***. In this case, the next P+ implant comprises just one step, for example using BF_2, with an energy of approximately 40 keV and a dose of $1*10^{14}$, given that it is no longer necessary to traverse the LDD region ***50***".

[55] This embodiment in practice enables just one implantation of the body-contact region ***61*** to be carried out, at the expense of more critical aspects linked to the LDD mask ***45***.

[56] In both cases, the formation of the body-contact regions ***61*** with open contacts and thus self-aligned to the respective contacts ***38*** enables a reduction in the distances between the various regions of the device, without introducing critical aspects; in particular, it is possible to reduce to 0.4 μm the distance between the body contacts 38 and the gate region ***11***. By further reducing the dimensions of the contacts and the distance between the edge of the drain contacts and the corresponding drain active area, it is possible to reduce the pitch of DMOS devices, using the indicated technology, from 4.1 to 3.3 μm. In practice, a 20% reduction in the pitch and area of the DMOS is obtained.

[57] DMOS transistor formed according to the described embodiments may be utilized in a variety of different types of electronic systems, such as computer systems.

[58] Finally, it is evident that modifications and variations may be made to the device and fabrication process described herein, without departing from the scope of the present invention. In particular, it is emphasized that the conductivity of the various regions may be opposite to the indicated, with a well of P type, a body region of N type, and source and drain regions of P type.

[59] From the foregoing it will be appreciated that, although specific embodiments of the invention have been described herein for purposes of illustration, various modifications may be made without deviating from the spirit and scope of the invention.